WHAT IS CLAIMED IS:

1. A probe designing method of designing a base sequence to be used as a probe which is hybridized with an unknown nucleic acid fragment to perform gene analysis, comprising:

the generation step of generating a tree in which a plurality of partial base sequences obtained on the basis of a target base sequence are arranged on nodes;

the evaluation step of evaluating the suitability as a probe of a partial base sequence represented by a desired node, on the basis of partial base sequences indicated by nodes present on that path on the tree, which is connected to the desired node; and

the determination step of determining a partial base sequence to be used as a probe on the basis of the evaluation result in the evaluation step.

2. The method according to claim 1, wherein the plurality of partial base sequences in the generation step are partial base sequences obtained from a complementary base sequence of the target base sequence.

3. The method according to claim 1, wherein

the plurality of partial base sequences in the generation step are partial base sequences obtained from the target base sequence, and

the determination step comprises selecting a partial base sequence on the basis of the evaluation result in the evaluation step, and determining a

EP 0 736 545 A2

complementary base sequence of the selected partial base sequence as a partial base sequence to be used as a probe.

4. The method according to claim 1, wherein the generation step comprises generating a tree for typing all partial base sequences obtained on the basis of the target base sequence.

5. The method according to claim 1, wherein the evaluation step comprises introducing an evaluation function which, when a base sequence whose specificity with respect to the target changes exists near the center of a partial base sequence represented by the desired node, evaluates that the partial base sequence is adequate as a probe.

6. The method according to claim 1, wherein the evaluation step comprises calculating the entropy of each node present on the path on the basis of the number of times of appearance, in the target, of a partial base sequence corresponding to the node, and performing evaluation on the basis of a decrease of the calculated entropy.

7. The method according to claim 6, wherein the evaluation step comprises introducing an evaluation function which attaches importance to a change in the entropy near the center of a partial base sequence corresponding to the desired node.

8. The method according to claim 5, wherein the

EP0845534A2

determination step comprises determining, as a probe, a partial base sequence corresponding to a node whose value calculated by the evaluation function in the evaluation step exceeds a predetermined value.

9. The method according to claim 6, wherein the determination step comprises determining, as a probe, a partial base sequence corresponding to a node whose change in the entropy exceeds a predetermined value.

10. The method according to claim 1, further comprising the grouping step of grouping the plurality of partial base sequences in accordance with specificity with respect to the target base sequence,

wherein the determination step comprises determining a partial base sequence to be used as a probe from each group on the basis of the evaluation result in the evaluation step.

11. The method according to claim 1, further comprising:

the grouping step of grouping the plurality of partial base sequences in accordance with specificity with respect to the target base sequence; and

the selecting step of selecting a group having specificity appropriate as a probe from groups obtained in the grouping step,

wherein the determination step comprises determining a partial base sequence to be used as a probe, from each group selected in the selecting step,

10081554.022202

on the basis of the evaluation result in the evaluation step.

12. The method according to claim 11, wherein the selecting step comprises selecting only a necessary and sufficient group completely independent in terms of information.

13. The method according to claim 11, wherein the selecting step comprises eliminating at least a group having no specificity with respect to all targets to be analyzed.

14. The method according to claim 11, wherein the evaluation step comprises evaluating partial base sequences in a group selected in the selecting step.

15. The method according to claim 10, wherein

the target contains a plurality of base sequence patterns, and

the grouping step comprises assigning, to the same group, partial base sequences which react or do not react identically with each of the plurality of base sequence patterns.

16. The method according to claim 1, wherein in the tree, the base sequence order of partial base sequences represented by node connections is consistent with the base sequence order in the target.

17. The method according to claim 1, wherein in the tree, the base sequence order of partial base sequences represented by node connections is changed such that the

2002-4554-0202-02

central one of corresponding partial base sequences in the target is the first one.

18. The method according to claim 1, wherein the evaluation step comprises evaluating only a partial base sequence having a length within a previously designated range.

19. The method according to claim 1, wherein the evaluation step comprises evaluating only a partial base sequence meeting a melting temperature condition within a previously designated range.

20. The method according to claim 1, wherein the determination step comprises determining a partial base sequence as a probe, from partial base sequences having lengths within a previously designated range, on the basis of the evaluation result in the evaluation step.

21. The method according to claim 1, wherein the determination step comprises determining a partial base sequence as a probe, from partial base sequences meeting a melting temperature condition within a previously designated range, on the basis of the evaluation result in the evaluation step.

22. A probe designing method of designing a base sequence to be used as a probe which is hybridized with an unknown nucleic acid fragment to perform gene analysis, comprising:

the generation step of generating a partial base sequence hash table for typing partial base sequences

obtained on the basis of a target base sequence and having a specific length;

the evaluation step of evaluating the suitability as a probe of a partial base sequence present in the base sequence hash table, on the basis of the base sequence thereof; and

the determination step of determining a partial base sequence to be used as a probe on the basis of the evaluation result in the evaluation step.

23. The method according to claim 22, wherein the partial base sequences in the generation step are partial base sequences obtained from a complementary base sequence of the target base sequence.

24. The method according to claim 22, wherein

the partial base sequences in the generation step are partial base sequences obtained from the target base sequence, and

the determination step comprises selecting a partial base sequence on the basis of the evaluation result in the evaluation step, and determining a complementary base sequence of the selected partial base sequence as a partial base sequence to be used as a probe.

25. The method according to claim 22, wherein the generation step comprises generating a plurality of hash tables in accordance with partial base sequences having different lengths.

26. The method according to claim 22, wherein the evaluation step comprises introducing an evaluation function which, when a base sequence whose specificity with respect to the target changes exists near the center of a partial base sequence, evaluates that the partial base sequence is adequate as a probe.

27. The method according to claim 22, wherein

a plurality of targets exist, and

the evaluation step comprises obtaining a specific position at which base sequences are different between a plurality of base sequences of the plurality of targets, and evaluating the suitability as a probe on the basis of the specific position in a partial base sequence registered in the hash table.

28. The method according to claim 27, wherein the evaluation step comprises checking whether the specific position is in the center of a base sequence, in order to evaluate the suitability as a probe.

29. The method according to claim 26, wherein the determination step comprises selecting a probe whose value calculated by the evaluation function in the evaluation step exceeds a predetermined value.

30. The method according to claim 22, further comprising the grouping step of grouping the plurality of partial base sequences in accordance with specificity with respect to the target,

wherein the determination step comprises

EP 2 525 304 A1

determining a partial base sequence to be used as a probe from each group on the basis of the evaluation result in the evaluation step.

31. The method according to claim 22, further comprising:

the grouping step of grouping the plurality of partial base sequences in accordance with specificity with respect to the target; and

the selecting step of selecting a group having specificity appropriate as a probe from groups obtained in the grouping step,

wherein the determination step comprises determining a partial base sequence to be used as a probe, from each group selected in the selecting step, on the basis of the evaluation result in the evaluation step.

32. The method according to claim 31, wherein the selecting step comprises selecting only a necessary and sufficient group completely independent in terms of information.

33. The method according to claim 31, wherein the selecting step comprises eliminating at least a group having no specificity with respect to a plurality of targets to be analyzed.

34. The method according to claim 31, wherein the evaluation step comprises evaluating partial base sequences in a group selected in the selecting step.

35. The method according to claim 30, wherein

the target contains a plurality of base sequence patterns, and

the grouping step comprises assigning, to the same group, partial base sequences which react or do not react identically with each of the plurality of base sequence patterns.

36. The method according to claim 22, wherein the evaluation step comprises evaluating only a partial base sequence meeting a melting temperature condition within a previously designated range.

37. The method according to claim 22, wherein the determination step comprises determining a partial base sequence as a probe, from partial base sequences meeting a melting temperature condition within a previously designated range, on the basis of the evaluation result in the evaluation step.

38. A probe designing method of designing a base sequence to be used as a probe which is hybridized with an unknown nucleic acid fragment to perform gene analysis, comprising:

the generation step of generating a discrimination tree for typing a list of a plurality of partial base sequences obtained from target base sequence data;

the evaluation step of evaluating the suitability as a probe of a probe candidate present in the discrimination tree; and

the selecting step of selecting a probe to be used on the basis of the evaluation result in the evaluation step.

39. The method according to claim 38, wherein the target base sequence data contains all base sequences which can exist in a specimen.

40. The method according to claim 38, wherein the target base sequence data contains all base sequences which can exist in a specimen and a specific base sequence.

41. The method according to claim 40, wherein the evaluation step comprises evaluating a probe candidate formed by a partial sequence of the specific base sequence.

42. A probe designing method of designing a base sequence to be used as a probe which is hybridized with an unknown nucleic acid fragment to perform gene analysis, comprising:

the generation step of generating a partial base sequence hash table for typing a list of a plurality of partial base sequences obtained from target base sequence data and having a specific length;

the evaluation step of evaluating the suitability as a probe of a probe candidate present in the partial base sequence hash table; and

the selecting step of selecting a probe to be used on the basis of the evaluation result in the evaluation

step.

43. The method according to claim 42, wherein the target base sequence data contains all base sequences which can exist in a specimen.

44. The method according to claim 42, wherein the target base sequence data contains all base sequences which can exist in a specimen and a specific base sequence.

45. The method according to claim 44, wherein the evaluation step comprises evaluating a probe candidate formed by a partial sequence of the specific base sequence.

46. An information processing apparatus for realizing the probe designing method according to claim 1.

47. A program for allowing a computer to realize the probe designing method according to claim 1.

48. A storage medium storing a program for allowing a computer to realize the probe designing method according to claim 1.

49. A DNA microarray comprising a base probe determined by using the probe designing method according to claim 1.

50. A gene inspecting apparatus comprising a base probe determined by using the probe designing method according to claim 1.

EP 1 564 659 A2